

2026 年（令和 8 年）4 月 3 日

約 12 兆トークンの良質なコーパスで学習した新たな国産 LLM 「LLM-jp-4 8B モデル」 「LLM-jp-4 32B-A3B モデル」を オープンソースライセンスで公開 ～一部ベンチマークで GPT-4o や Qwen3-8B を上回る性能を達成～

大学共同利用機関法人 情報・システム研究機構 国立情報学研究所（NII、所長：黒橋 禎夫、^{くろはし きだお}東京都千代田区）の大規模言語モデル研究開発センター（LLMC）は、同センターが主宰する LLM 研究開発コミュニティ「LLM-jp」の活動の中で大規模言語モデル（LLM）のフルスクラッチ学習を実施し、約 86 億パラメータ^{(*)1}の「LLM-jp-4 8B モデル」と約 320 億パラメータの MoE モデル^{(*)2}「LLM-jp-4 32B-A3B モデル」をオープンソースライセンスで一般公開しました。公開モデルの学習では、オープンソース AI の定義（OSAID）に配慮し、第三者も入手可能な良質な学習コーパス^{(*)3}の収集・選別・構築を行い、インターネット上の公開データや政府・国会の文書、合成データなどからなる約 12 兆トークンの学習コーパスを整備・使用しました。公開モデルは最大で約 6 万 5 千トークンの入出力まで処理でき、言語モデルの日本語理解能力を測る「日本語 MT-Bench」、英語理解能力を測る「MT-Bench」において、強力な多言語 LLM である「GPT-4o」や「Qwen3-8B」を上回る性能を達成しています。

LLMC では「LLM-jp-4 8B モデル」と MoE モデル「LLM-jp-4 32B-A3B モデル」を活用して LLM の透明性・信頼性の確保に向けた研究開発を進めていきます。また、現在、より大規模なパラメータを備えたモデルの開発を進めており、2026 年度に順次公開予定です。

今回公開した LLM の概要

(1) 利用計算資源

国立研究開発法人産業技術総合研究所が提供する AI 橋渡しクラウド (ABCI 3.0) を利用しました。

(2) モデル

- LLM-jp-4 8B モデル
 - モデルアーキテクチャ : Llama 2 アーキテクチャ
 - パラメータ数 : 約 86 億個 (8B)
- LLM-jp-4 32B-A3B モデル
 - モデルアーキテクチャ : Qwen3 MoE アーキテクチャ
 - 総パラメータ数 : 約 320 億個 (32B)
 - 総エキスパート数 : 128
 - アクティブパラメータ数 : 約 38 億 (3.8B)
 - アクティブエキスパート数 : 8

(3) 学習コーパス

オープンソース AI の定義 (OSAID) に配慮し、第三者も入手可能な良質な学習コーパスの収集・選別・構築を行った上で、LLM-jp が過去に開発・公開した「LLM-jp-3.1」シリーズと比較して約 6 倍の規模となる学習コーパスを構築しました。

事前学習

事前学習では、インターネット上の公開データ、政府・国会の文書などから構成される大規模な事前学習コーパスを使用しました。本コーパスは、総計約 19.5 兆トークンであり、日本語約 7,000 億トークン、英語約 17.8 兆トークン、他言語 (中国語・韓国語) 約 8,500 億トークン、プログラムコード約 2,000 億トークンで構成されています。各サブコーパスをどの程度重点的に学習に使用するかを実験を通して最適化し、合計約 10.5 兆トークンを事前学習に使用しました。

中間学習

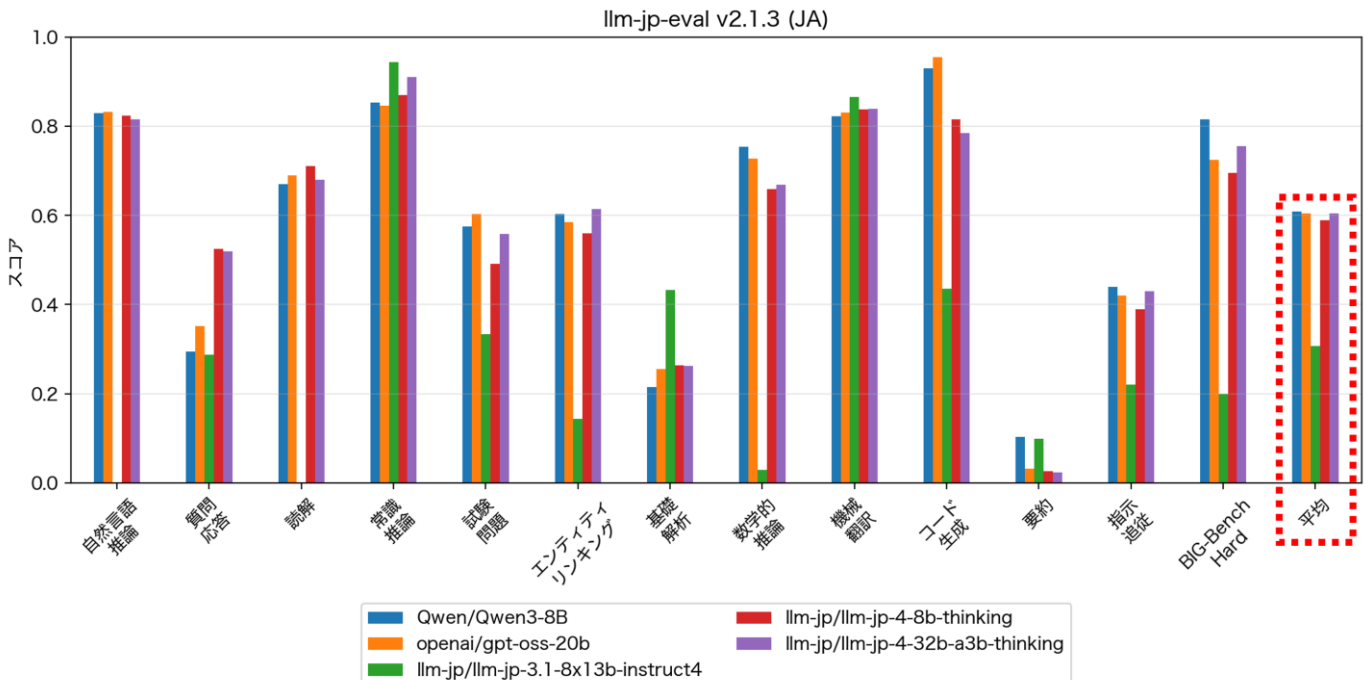
事前学習に続き中間学習を実施しました。中間学習には、事前学習コーパスに指示事前学習データ (Instruction Pre-training データ) を含む LLM による合成データを加えた、合計 1.2 兆トークンの学習コーパスを使用しました。

(4) チューニング

英語および日本語インストラクションチューニングデータ 22 種類を用いてチューニングを実施しました。学習データは、オープンソースライセンスのデータのほか、LLM-jp が開発したデータを含みます。LLM-jp にて開発したデータは順次公開予定です。

(5) 評価

- LLM-jp が開発する評価フレームワーク「llm-jp-judge」を用い、GPT-5.4 による LLM-as-a-Judge 評価を行いました。日本語理解性能を測る「日本語 MT-Bench」では、「LLM-jp-4 8B モデル」が 7.54、「LLM-jp-4 32B-A3B モデル」が 7.82 のスコアを達成しました。これは、強力な多言語 LLM である「GPT-4o」の 7.29 や「gpt-oss-20b」の 7.33、「Qwen3-8B」の 7.14 を上回るものです。また、英語理解性能を測る「MT-Bench」では、「LLM-jp-4 8B モデル」が 7.79、「LLM-jp-4 32B-A3B モデル」が 7.86 を達成しました。これは、「GPT-4o」の 7.69、「gpt-oss-20b」の 7.85、「Qwen3-8B」の 7.69 と比較して、同等以上の性能です。
- LLM-jp が開発する既存の日本語・英語の言語資源に基づく 42 種類の評価データを用いて横断的な評価を行うフレームワーク「llm-jp-eval v2.1.3」を用いた評価を実施し、「LLM-jp-4 8B モデル」、「LLM-jp-4 32B-A3B モデル」ともに日本語性能において「gpt-oss-20b」、「Qwen3-8B」と同等の性能を達成することを確認しました。



代表的な LLM の llm-jp-eval による項目別評価

(6) 開発モデル・ツール・コーパスの公開 URL

<https://llm-jp.nii.ac.jp/release>

今後の予定

本モデルの成果を踏まえ、より大規模なパラメータを備えた「LLM-jp-4 32B モデル」と MoE モデル「LLM-jp-4 332B-A31B モデル」の開発を進めています。あわせて、実運用での利便性を考慮した軽量モデルの開発も並行して進めます。これらのモデルは、2026 年度中に順次公開予定です。

(参考 1) LLM-jp の概要

1. NII が主宰する LLM-jp では、自然言語処理及び計算機システムの研究者を中心として、大学・企業等から 2,600 名以上（2026 年 3 月 31 日現在）が集まり、ハイブリッド会議、オンライン会議、Slack 等を活用して LLM の研究開発について情報共有を行うとともに、共同で LLM 構築等の研究開発を行っています。具体的には、以下の目的で活動しています。
 - オープンかつ日本語に強い LLM の構築とそれに関連する研究開発の推進
 - 上記に関心のある自然言語処理および関連分野の研究者によるモデル構築の知見や最近の研究の発展についての定期的な情報交換
 - データ・計算資源等の共有を前提とした組織横断的な研究者間の連携の促進
 - モデル・ツール・技術資料等の成果物の公開
2. 「コーパス構築 WG」「モデル構築 WG」「チューニング・評価 WG」「マルチモーダル WG」「実環境インタラクション WG」「対話 WG」「学術ドメイン WG」「安全性 WG」「原理解明 WG」を設置し、それぞれ、早稲田大学 河原大輔教授、東北大学 鈴木潤教授、東京大学 宮尾祐介教授、東京科学大学 岡崎直観教授、早稲田大学 尾形哲也教授、名古屋大学 東中竜一郎教授、国立情報学研究所 相澤彰子教授、国立情報学研究所 関根聡特任教授、東京大学 大関洋平准教授を中心に研究開発活動に取り組んでいます。このほか、東京大学 田浦健次朗教授、東京科学大学 横田理央教授（並列計算手法等）を筆頭に、多数の方々の貢献により活動を進めています。
3. 詳細については、ホームページ <https://llm-jp.nii.ac.jp/> をご参照ください。

(参考 2)

本成果は文部科学省補助金事業「生成 AI モデルの透明性・信頼性の確保に向けた研究開発拠点形成」の結果得られたものです。

謝辞

本研究は、国立研究開発法人産業技術総合研究所および株式会社 AIST Solutions が提供する ABCI 3.0「開発加速利用」の支援を受けて実施しました。

本モデルの開発にあたり、「生成 AI モデルに資する言語資源の研究開発」に関する共同研究契約に基づき、国立国語研究所（NINJAL）より国語研日本語ウェブコーパス（whole-NWJC）の提供を受けました。また「国立国会図書館及び学術情報センターの相互協力に関する協定」に基づく協力の一環として、国立国会図書館（NDL）より、国立国会図書館インターネット資料収集保存事業（WARP）で収集された Web サイトの URL リストの提供を受けました。

〈メディアの皆様からのお問い合わせ先〉

大学共同利用機関法人 情報・システム研究機構 国立情報学研究所
総務部総務企画課 企画・広報チーム

TEL:03-4212-2164 E-mail : media@nii.ac.jp

(*1) **パラメータ数** : 大規模言語モデルは言語を学習した大規模なニューラルネットワークで、パラメータはニューラルネットワークの規模を示す指標のひとつ。パラメータ数が多いほど高い性能であるといわれている。

(*2) **MoE (Mixture of Experts) モデル** : LLM の内部に複数のエキスパート（専門家）を内包しているモデルで、推論時にそれらのエキスパートを動的に切り替えることで、効率的な推論を可能にしているモデル。

(*3) **コーパス** : 自然言語の文章を構造化し大規模に集積したデータベース。